THE COCKPIT COGNITIVE WALKTHROUGH

Peter G. Polson
University of Colorado
Boulder, CO 80309

Nancy Smith
NASA-Ames Research Center
Moffett Field, CA 94035-1000

ABSTRACT

The Cockpit Cognitive Walkthrough is a usability inspection method that evaluates interactions between a cockpit procedure and an avionics interface by showing that the interface supports execution of the procedure by providing feedback for correct pilot actions and error recovery and by guiding the execution of a novel or an infrequently performed procedure. This paper describes the method and summarizes two evaluation studies current in progress.

INTRODUCTION

The Cockpit Cognitive Walkthrough is a new design and evaluation method for prototype interfaces, cockpit procedures, and training materials for glass cockpit aircraft. It is a usability inspection method (Nielsen & Mack, 1994) that evaluates interactions between a cockpit procedure and an avionics interface. The method evaluates the usability of the interface by showing that it supports execution of the procedure by providing feedback for correct pilot actions and error recovery and by guiding the execution of a novel or an infrequently performed procedure. This focus on providing adequate support for exploration also improves other attributes of usability, including ease of learning and ease of use.

The Cockpit Cognitive Walkthrough is based on the cognitive walkthrough (Wharton, Lewis, Rieman, and Polson, 1994). The cognitive walkthrough evaluates the ability of a skilled user of an environment like the MAC OS to perform novel or occasionally performed tasks by exploration. The evaluation criteria are derived from a theory of performing a task by exploration (Kitajima and Polson, 1997) and empirical evidence supporting the theory.

Problem

Recent articles on the operation of glass cockpit aircraft have documented various operational and training issues (e.g., Billings, 1997; Sarter & Woods, 1995; Sherry & Polson, 1999). Most researchers have concluded that these problems are caused by interactions between the underlying complexity of the avionics and limitations of current avionics interfaces and of pilot training programs. Various investigators have discussed the problems with avionics displays and the quality of feedback provided to pilots in glass cockpits (Hutchins, 1996; Sherry & Polson, 1999).

There are numerous model-based and empirically based evaluation methods that have been developed by researchers in human factors in computer systems (Nielsen & Mack, 1994). However, airframers and airlines do not systematically utilize these methods. Many of these methods are extremely time consuming and require skills not normally available in these organizations (e.g. writing complex cognitive simulation programs). Airlines and airframers do make extensive use of formal design reviews that include pilot participants in developing and evaluating designs, but details of these review processes and their evaluation criteria are not explicitly defined.

Solution

The Cockpit Cognitive Walkthrough is a human factors evaluation method that provides a set of criteria for evaluating design alternatives. It should be usable in these organizations because it does not require radical changes to the processes they currently use to develop avionics interfaces, cockpit procedures, and training materials.

Performing by exploration. The Walkthrough assumes that designing avionics interfaces to support performing by exploration will solve many of the human factor problems with modern avionics that were described in the preceding section.

Performing by exploration is a problem solving process guided by knowledge of the task to be performed, of how to execute related tasks, and knowledge of the task interface conventions. Optimizing an interface to support performing by exploration also optimizes other important attributes of usability including ease of learning and ease of use. Consistency of both the properties of interfaces and the structure of procedures are critical for supporting exploration. Explicit and enforced interface conventions enable a user to infer correct actions for novel controls and displays. Consistent structures for procedures enable users to generalize across related tasks.

Many tasks in the cockpit are learned or performed by exploration. A very small fraction of the functionality of the avionics is explicitly trained. Pilots' learn by exploration on the line. In addition, many tasks are performed infrequently or are novel versions of trained tasks. These tasks are also performed by exploration. Thus support for exploration is a necessary feature of a usable avionics interface or effective cockpit procedure.

Labeling and feedback. These two terms refer to labels on controls, CDU page and line labels, and information presented on other cockpit displays. Labeling and feedback have four different functions. The first function is to support formulation of correct goals or task descriptions. This function is critical for both skilled performance as well as performing by exploration. The second function is guiding action selection. When performing by exploration, pilots use control labels and other cues from a system interface to guide selection of a correct action. The third function of feedback is to signal the consequences of the last action and to enable a pilot to decide whether or not to continue. The fourth function of feedback is enabling the pilot to deal with interruptions. Superior labeling and feedback provide pilot with unambiguous information about the current state of the interrupted task. When the interrupted task is resumed, the pilot can rapidly reconstruct the task state and successfully complete the task.

DESCRIPTION OF THE COCKPIT COGNITIVE WALKTHROUGH

The walkthrough takes as input the pilots' descriptions of a task, the sequence of actions required to perform the task, and the display changes that result from each action. The method evaluates the ability of a pilot to perform the task by exploration.

This analysis should be performed as early as possible in the development cycle of a new system. The goal is to identify usability problems before completing the detailed design of the system.

To prepare for the Walkthrough, the design should be documented as a series of "storyboards" showing the system interface at each step of each task. In addition, each storyboard should describe the action or actions necessary to complete each step and the feedback provided by the interface during the execution of each step. Note that since detailed task and interface descriptions are important inputs to the design process, they are often already available early in development, so it may require very little additional effort on the part of system designers to create the storyboard needed for the Walkthrough.

Goals and Subgoals.

The Cockpit Cognitive Walkthrough assumes that performing by exploration is guided by a pilot's representation of the task to be performed. This representation is in the form of a goal to perform the task and a collection of subgoals that describes each of the major subtasks. A pilot's representation of a task goal will be determined by how the task is initiated and presented to him. In addition, the initiating event will determine the context in which the task must be performed, including time pressure and workload. A task may be initiate by a checklist item, an SOP retrieved from memory, an ATC directive, an alert or warning, or an abnormal condition detected by pilot. *The exact phraseology used in the initiating event will have a strong influence on a pilot's characterization of the goal.*

Subgoal structures. A task of any complexity (more than three or four steps) will be decomposed into a series of subtask each represented by a subgoal. The subgoals will be determined by interactions between the substructure of the original task and the details of the system interface. Complying with a multi-element clearance is a good example of a task whose subgoals are determined by the subcomponents of the initiating event.

Supporting performing by exploration of tasks with complex subgoal structures puts real demands on the quality of the feedback provided by the interface. The feedback must support the initiation and correct formulation of each subgoal. In addition, feedback must support recognition of subtask completion and termination of an associated subgoal.

Characterization of goals and subgoals is the most difficult and critical part of an Cockpit Cognitive Walkthrough. However, the aviation environment does simplify this process. Regulations, airline standard operating procedures, and the air traffic control system define and characterize many of the tasks performed by pilots. There are agreed upon descriptions of these tasks and specified conditions for their performance.

Doing the Walkthrough

Doing the Walkthrough involves evaluating each step portrayed in the storyboard for a task. The Walkthrough is a hand simulation of the processes that support successful performing by exploration. Related models of performing by exploration have been transformed into a series of questions that the analyst asks about each step portrayed in the storyboard. The questions, taken from Wharton, Rieman, Lewis, & Polson (1994) and adapted for an aviation environment, are:

- Will the pilot be trying to do the right thing?
- Will the pilot know that the correct action is available?

- Will the pilot connect the correct action to what they are trying to do?
- If the correct action is made, will the pilot see that things are going OK? (And if an *incorrect* action is performed, will feedback be sufficient for the pilot to detect the error?)

In the follow sections, we describe the details of evaluating the answers to each of these questions and present the rationale for each derived from the Kitajima and Polson (1997) framework.

Will the pilot be trying to do the right thing? In attempting to answer this question, the group performing the walkthrough should try to determine how the pilot goes about generating the exact wording of the goal or subgoal. Possible answers include the reading of a checklist item, correct interpretation of an ATC clearance or feedback generated by the avionics. The goal or subgoal can also be generated by information stored in memory.

Will the pilot know that the correct action is available? In order to select an action, the pilot has to know that it is possible. This may sometimes become an issue in the cockpit. In the process of adding new functions to the FMS that are accessed through the CDU, for example, designers may incorporate new items on the pages accessed by one of the page-mode keys (e.g., MENU). If pilots do not know about the new function, they will be unlikely to press the MENU key to access it.

Will the pilot connect the correct action to what they are trying to do? There is a large amount of evidence in the human-computer interaction literature that users performing by exploration will select actions (e.g., pressing a page-mode key) whose descriptions are similar to the user's current goal (Kitajima & Polson, 1997; Wharton, et al., 1994). The same tendencies have been found in the cockpit with pilots who are being training to program the FMS (Polson, Irving, & Irving, 1995).

Polson and Lewis (1990) called this tendency the "label-following strategy". The label following strategy guides action planning in performing by exploration. A yes response to "Will the pilot connect the correct action to what they are trying to do?" means that the label-following strategy will work. The primary focus of the Cockpit Cognitive Walkthrough is to evaluate the effectiveness of the label-following strategy for a proposed interface design or new cockpit procedure.

If the correct action is made, will the pilot see that things are going OK? Answering this question involves evaluating the adequacy of feedback provided by the interface. Performing by exploration is a problem-solving task. Pilots will have varying degrees of certainty about the correctness of each of their actions. Good feedback confirms correct actions or provides the information needed for error correction.

Specializing The Walkthrough for Tasks Using the CDU

The main uses of the walkthrough at NASA Ames Research Center, summarized below, have been to evaluate interfaces to new functionality to the FMS, in particular use of datalink for ATC communications. These design problems were constrained by the existing hardware for the CDU interface and current cockpit displays. The only degrees of freedom in these designs were the organization of new pages implementing the new functionality and the page titles (labels) and line labels. Recommendations for changes in a design were confined to changes to CDU page content and organization and CDU page and line labels.

In addition, it was assumed that pilots performing tasks involving these new functions would have extensive experience with the FMS and knowledge of basic CDU interface conventions.

The ***ACCESS***, ***DESIGNATE***, and ***INSERT*** Templates

Polson, Irving, and Irving (1995) developed GOMS models for a set of flight planning and preflight initialization tasks that use the CDU to program the FMS in the Boeing 737-300/400/500. They identified generic goal structures that describe three functional subgroups of keystroke sequences that were common to many of these basic CDU tasks. These three subgoal templates are **ACCESS**, **DESIGNATE**, and **INSERT**.

ACCESS involves two sub-subgoals: (1) identifying the CDU page associated with the current task, then (2) carrying out the sequence of operations necessary to get that page displayed on the CDU. The **DESIGNATE** subgoal is invoked any time the pilot must formulate and enter information into the CDU. The **INSERT** subtask involves placing the designated information into the appropriate line field(s) on the currently displayed CDU page.

Execution of the **ACCESS**, **DESIGNATE**, and **INSERT** subgoals is guided by the description of the task being performed using the CDU and knowledge of these templates. The walkthrough focuses on showing how the task description (task goal) guides the accomplishment of each of these subgoals. In the discussion that follows, the description of the task goal is taken as a given. In a complete walkthrough, the analyst would have to consider how the goal was initiated and described.

Performing **ACCESS**. **ACCESS** can be a problematic subgoal because the designer can't modify the labels on the most powerful action for **ACCESS** available on the CDU, the page/mode keys. A new data link function can be added to the list of functions on the MENU page accessed by the MENU key or by the ATC key. However, such generic key labels are not good matches to any task description. Thus, pilots will have to just remember the initial step of most new functions added to the CDU interface. This is a serious flaw in an interface to any infrequently performed task.

However, once the pilot has accessed the top-level page of the collection of new pages that implement the new functionality, the designer has complete control over the page titles and the labels on the line select keys that bring up other pages in the collection. The major focus of the walkthrough of the **ACCESS** function is to ensure that these labels provide the guidance (label following) necessary to access the correct page.

Page titles are also critical in providing feedback and supporting error recovery. A page title that matches the task description confirms the pilot's last **ACCESS** action. A mismatching title should initiate error recovery in a well-designed collection of pages.

Performing **DESIGNATE**. **DESIGNATE** is the most heterogeneous of the subgoals. It can range from a single line select action to a complex sequence involving **ACCESS** to other CDU pages, scanning other cockpit displays, accessing other flight documents, or transforming an ATC clearance into a sequence of flight plan modifications.

Typical examples of complex **DESIGNATE** subtasks are vertical flight plan modifications where several parameters have to be entered into the scratch pad of the CDU. Relevant information may have to be retrieved from memory or come from other CDU pages, an ATC clearance, the flight plan, or the FCOM. It can be difficult to support performing by exploration for such complex sequences of cognitive and actions given the limitations of the CDU interface and the possible need to acquire information from multiple sources.

Performing **INSERT**. **INSERT** involves transferring information entered into the scratch pad during the last **DESIGNATE** subtask into the appropriate line field on the currently displayed CDU page. Good matches between the correct line label and pilot's current goal guide this subtask.

EVALUATION OF THE COCKPIT COGNITIVE WALKTHROUGH

The next section will provide an illustration of the Cockpit Cognitive Walkthrough method by describing its use in two different projects at NASA Ames Research Center. In the first project, the Walkthrough was used to refine the design of a CDU data link interface in preparation for a full mission simulation experiment. The second project used the Walkthrough to evaluate the flight deck interface for the FANS-1 controller-pilot data link communications system (CPDLC) (Smith, et al, 1999).

Example 1: Using the Walkthrough for Interface Refinement

One research group at NASA is exploring means to improve the compatibility of flight deck automation with ground-based decision support tools for air traffic control (the Center TRACON Automation System, or CTAS). An experiment is planned that will explore the use of data link for ground-to-air uplink of CTAS-generated route clearances in the terminal approach control area. The initial design for the flight deck data link interface for this experiment was based on the CDU pages that support FANS-1 CPDLC in the oceanic airspace. The Walkthrough was used to refine this initial design and improve its suitability for the experimental task.

Method The test plan for this experiment included detailed descriptions of pilot actions, data link messages, FMS operations, CDU pages, and CTAS operations for each step in the scenario. This scenario description was easily translated into storyboard representation using PowerPoint slides to represent the sequence of CDU page changes that preceded and followed each pilot action (a sample slide from the storyboard is shown in Figure 1). Pilot activities associated with receipt of the CTAS route uplink clearance included: (1) detecting the ATC uplink message, (2) accessing the message, (3) loading the uplinked route clearance into the FMS, (4) reviewing the modified route, (5) accepting (or rejecting) the uplinked clearance, and (6) executing the route modification. Figure 1 shows two of the slides developed for the Walkthrough analysis of this initial design.

During the session, the project team used the slides to walk through the task, step by step, and decide whether the interface appeared to provided adequate support to ensure successful performance of each action. Team members discussed whether there was a reasonable

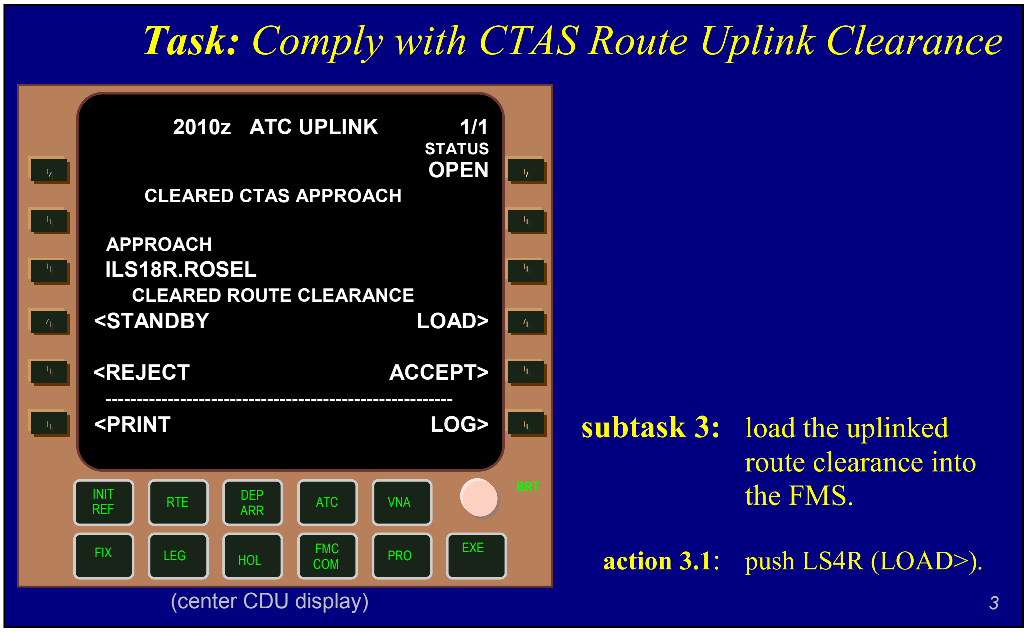

Figure 1. Slide #3 from storyboard for "Comply with CTAS Route Uplink Clearance."

The next correct action is to load the route uplink (cued by the "LOAD>" label). Accepting the clearance before loading and reviewing the route uplink is incorrect, so the "ACCEPT>" prompt should not be available at this stage in the task.

match between the pilot's probable intentions at each step and the label or cue that prompted the appropriate action.

Equally important was the identification of labels that could potentially mislead the pilot into performing the wrong action, or performing actions in the wrong order. Finally, system feedback that followed each action was assessed to determine whether it provided enough information for pilots to evaluate the correctness of the action performed and whether the task was completed. Suggestions for interface design changes were recorded during each session. Afterwards, a new set of slides was created, incorporating session material into a modified design that could be reviewed in a subsequent walkthrough.

This iterative process was repeated three times, with three one hour Walkthrough sessions conducted at roughly one week intervals. Examples of the interface changes that resulted from this design activity can be seen in Figure 2.

<u>Some observations</u> One rather striking observation was how little effort was required to conduct the walkthroughs and to implement the consequent interface changes. Because the task analysis had already been performed and interface mockups created as part of the experiment scenario development, creating the initial walkthrough slides was just a matter of translating the needed material from one format to another. One of the re-design goals was to try to keep associated software changes as minimal as possible--and indeed, one programmer on the project reported that it only took him about 20 minutes to implement the CCW recommendations. One interesting sidenote: project team members were located in three different parts of the country, and all three Walkthrough sessions were conducted by telecon. Group members each received a set of PowerPoint slides by e-mail before each session and referred to their own computers or printouts while conducting the walkthrough over the telephone.

Example 2: Using the Walkthrough to Evaluate an Operational System

The Walkthrough was also used as part of a NASA Ames investigation of human factors problems encountered by flight crews using FANS-1 CPDLC in the oceanic airspace. This project is described in more detail in a companion paper (Smith, et al., 1999). The Walkthrough was used to identify aspects of the FANS-1 task environment that might lead to performance problems. A survey based on this evaluation was developed then distributed to pilots that use FANS-1 CPDLC.

<u>Test of the CCW method.</u> Survey data will be used not only to the operational performance of FANS-1 CPDLC, but to validate the cockpit cognitive walkthrough: pilots' reports of operational problems will be compared with difficulties predicted by the walkthrough.

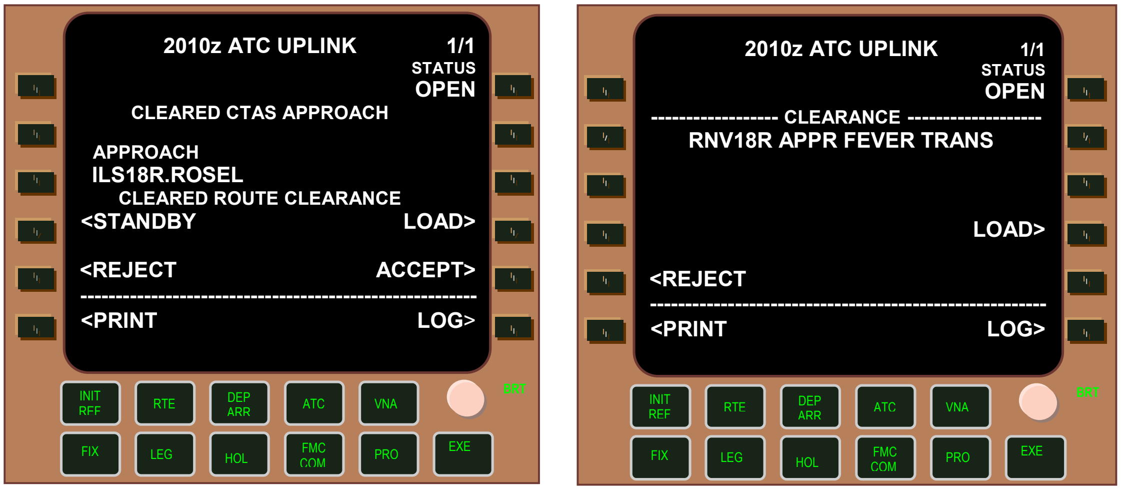

Figure 2. Initial draft *(left)* and final draft *(right)* of the CDU page that presents the ATC Uplink message in the NASA experiment. Changes include:

(1) inappropriate prompts removed ("ACCEPT>" and "<STANDBY").
(2) clearance label simplified to "---CLEARANCE---".
(3) clearance description matches title of referenced procedure chart ("RNV18R APPR FEVER TRANS").

CONCLUSIONS

The Cognitive Walkthrough is a usability inspection method that was developed to provide a practical procedure for evaluating the usability of office automation applications and other computer-based consumer production. The Cockpit Cognitive Walkthrough extends the use of this method to the flight decks of modern highly automated airliners. Wharton, et al. (1994) and Kitajima and Polson (1997) present the methods' theoretical foundations and supporting empirical studies in the office automation domain. In this paper, we have show that the original walkthrough can be extended to the cockpit and that the method generates explicit guidance to developers with a modest amount of effort. The method is practical because is adds a very structured evaluation step to the normal development process for avionics interfaces. The two evaluation studies in progress will evaluate the usability prediction made by the method. The method has already been extensively tested in the office automation domain and found to be practical and effective.

REFERENCES

Billings, C. E. (1997). Aviation automation: The search for a human-centered approach. Mahwhad, NJ: Lawrence Erlbaum.

Hutchins, E. (1996). *Immediate mode management interface final report*. NASA Ames Research Center.

Kitajima, M., & Polson, P.G. (1997). A comprehension-based model of exploration. Human-Computer Interaction.

Nielsen, J. & Mack, R. (1994). *Usability Inspection Methods*. New York: John Wiley & Sons.

Polson, P.G., Irving, S., and Irving, J.E. (1994) Final Report: Applications of Formal Models of Human-Computer Interaction to Training and Use of the Control and Display Unit. System Technology Division, ARD 200, Dept. of Transportation, Federal Aviation Administration. Also published as *ICS Technical Report*. 94-08. Boulder, CO: Institute of Cognitive Science, University of Colorado.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode errors and awareness in supervisory control. *Human Factors*, 37, 5–20.

Sherry, L., & Polson, P. G. (1999). Shared Models of FMS Vertical Guidance. *International Journal of Aviation Psychology*

Smith, N., Moses, J., Polson, P., Romahn, S., Palmer, E. (1999) A Survey of Flight Crew Experience with FANS-1 ATC Datalink. 10th International Symposium on Aviation Psychology.

Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994) The Cognitive Walkthrough Method: A Practitioner's Guide. In J. Nielsen and R. Mack (Eds.). *Usability Inspection Methods*. New York: John Wiley & Sons.